PERIODICITY AND APERIODICITY IN THE PERCEPTION OF SPEECH IN BOTH STEADY-STATE AND FLUCTUATING MASKERS

Kurt Steinmetzger & Stuart Rosen
UCL Speech, Hearing and Phonetic Sciences, London, UK

0344-001209

Introduction

Hearing-impaired listeners and cochlear implant (CI) users experience great difficulties in **speech** perception in all types of **background noise**, and unlike normal-hearing listeners show little benefit from fluctuations in the masker [1, 2].

One popular (if partial) explanation for these difficulties proposes a key role for **temporal fine structure (TFS)** cues [3]. However, there is controversy over whether TFS has a special role in allowing **fluctuating-masker benefit (FMB)** or whether its contribution to speech perception is just as important for steady maskers [4].

We investigated the abilities of normal-hearing listeners to perceive speech targets in the background of noise maskers in a variety of conditions **mixing presence and absence of periodicity** in both target and masker.

Methods

Speech Reception Thresholds **(SRTs)** were measured adaptively for **IEEE sentences** processed to change their source characteristics (and hence their periodicity).

Targets were produced either with a **Channel vocoder** [5] or **Tandem-Straight** [6] to produce speech with varying amounts of spectral detail.

T**he spectral resolution** of the target speech was varied over a wide range (7–24 channels).

In two separate experiments 16 and 12 normal-hearing subjects were tested.

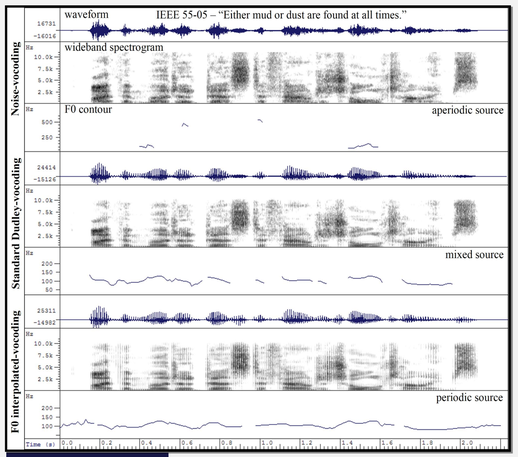

Figure 1. Targets

3 classes of target speech

- Noise-vocoded (fully aperiodic)
- Standard-vocoded (natural mixed source)
- F0 interpolated-vocoding (fully periodic)

4 different maskers

- Steady speech-shaped noise
- Amplitude-modulated speech noise (10 Hz sinusoidal)
- Harmonic complexes (with dynamically varying F0 contours)
- Amplitude-modulated harmonic complexes (10 Hz sinusoidal)

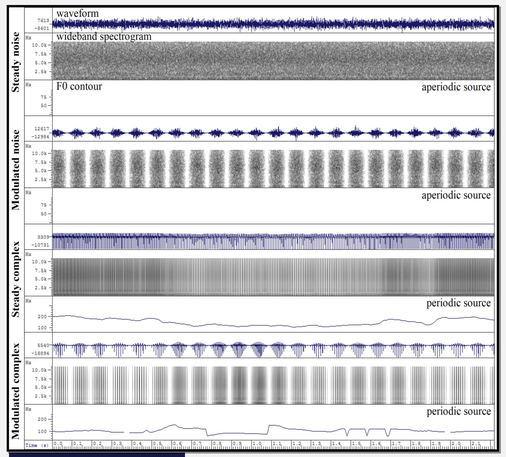

Figure 2. Maskers

Results

1) Speech Reception Threshold

More spectral detail leads to lower SRTs and thus more intelligible speech, with unprocessed speech scoring lowest in all 4 maskers.

Increasing **periodicity in the targets does not affect the SRTs**, supporting the notion of masker cancellation [7, 8].

Masker periodicity seems more important than masker fluctuations, as suggested by the stepwise pattern found for the 4 different maskers.

2) Fluctuating-masker Benefit

(SRT-difference of steady and fluctuating maskers)

Contrary to previous beliefs, the FMB is markedly **smaller for periodic than for aperiodic maskers**.

A substantial **FMB required a rather high spectral resolution** of the target speech (24 bands), while targets with low spectral detail (7 bands) led to **fluctuating-masker interference**.

A **natural mix of periodicity and aperiodicity** in the target speech was observed to foster the ability to glimpse, no matter if the masker was periodic or aperiodic (effects of up to **5 dB**).

3) Periodicity Benefit

(SRT-difference of aperiodic and periodic maskers)

Increasing **periodicity in the masker strongly aids speech intelligibility** across all levels of spectral resolution (effects of more than **10 dB**).

Periodicity Benefit was found for all 3 classes of target speech, so **does not depend on the amount of target periodicity**.

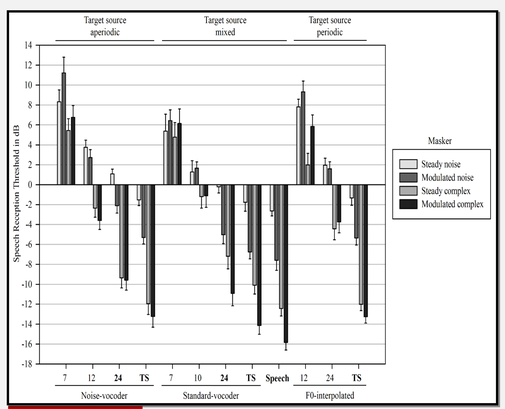

Figure 3. SRT

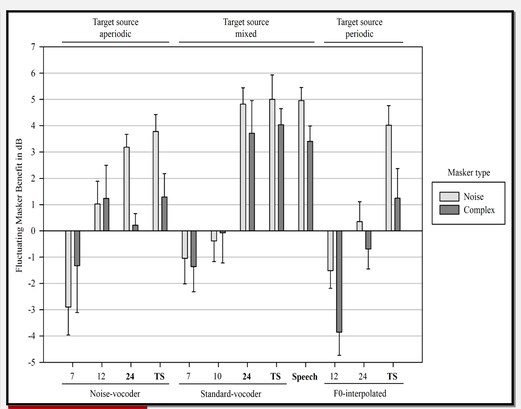

Figure 4. FMB

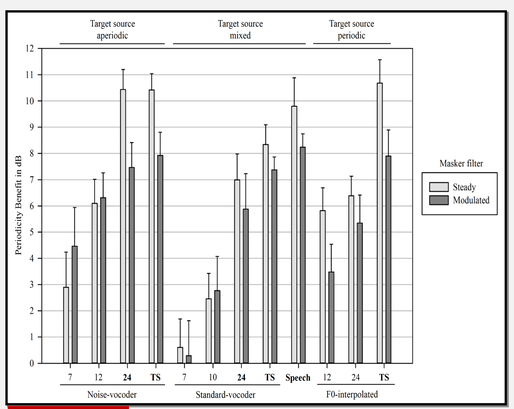

Figure 5. PB

Discussion

Our results show that **periodicity in the masker, but not the target speech,** greatly helps in tracking a speech signal through a background noise and suggest that the inability to exploit periodicity may be an even more important factor in the limitations of hearing-impaired and CI speech perception than the inability to benefit from fluctuations in a masker.

Masker periodicity had a **larger effect in steady-state maskers** (~10.5 dB) than in those that fluctuate in amplitude (~8 dB). Thus, it appears unlikely that sensitivity to TFS has a special role in glimpsing.

Regarding the ability to benefit from masker fluctuations the results of the present study indicate that it may not be the presence of TFS in the target speech as such but rather a **natural mix of periodicity and aperiodicity that promotes glimpsing**.

References

[1] Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *JASA, 88.*

[2] Nelson, P. B., Jin, S.-H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: cochlear implant users and normal-hearing listeners. *JASA, 113.*

[3] Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *PNAS, 103.*

[4] Hopkins, K., & Moore, B. (2008). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *JASA,125.*

[5] Dudley, H. (1939). Remaking speech. *JASA,11.*

[6] Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T., & Banno, H. (2008). Tandem-Straight: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation. *ICASSP.*

[7] de Cheveigné, A., McAdams, S., Laroche, J., & Rosenberg, M. (1995). Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement. *JASA, 97.*

[8] Vestergaard, M. D., & Patterson, R. D. (2009). Effects of voicing in the recognition of concurrent syllables. *JASA, 126.*

Acknowledgements

This project has been funded with support from the European Commission under Contract FP7-PEOPLE-2011-290000 and the Medical Research Council (Grant Number G1001255). We thank Hideki Kawahara and Jeanne Clarke for helpful advice and code concerning Tandem-Straight and Tim Green for helpful comments.